[Task1](#)

input file

Samsung|Optima|14|Madhya Pradesh|132401|14200
Onida|Lucid|18|Uttar Pradesh|232401|16200
Akai|Decent|16|Kerala|922401|12200
Lava|Attention|20|Assam|454601|24200
Zen|Super|14|Maharashtra|619082|9200
Samsung|Optima|14|Madhya Pradesh|132401|14200
Onida|Lucid|18|Uttar Pradesh|232401|16200
Onida|Decent|14|Uttar Pradesh|232401|16200
Onida|NA|16|Kerala|922401|12200
Lava|Attention|20|Assam|454601|24200
Zen|Super|14|Maharashtra|619082|9200
Samsung|Optima|14|Madhya Pradesh|132401|14200
NA|Lucid|18|Uttar Pradesh|232401|16200
Samsung|Decent|16|Kerala|922401|12200
Lava|Attention|20|Assam|454601|24200
Samsung|Super|14|Maharashtra|619082|9200
Samsung|Super|14|Maharashtra|619082|9200
Samsung|Super|14|Maharashtra|619082|9200

Output file

acadgild@localhost ~]$ hadoop jar tvinvalid.jar com.acadgild.hadoop.mapreduce.tvsales.invalidtv /
hadoop1/mapreduce/television.txt [/hadoop1/mapreduce/tvinvalidout](#)
18/06/19 19:35:44 WARN util.NativeCodeLoader: Unable to load native-hadoop library for your
platform... using builtin-java classes where applicable
18/06/19 19:35:45 INFO client.RMProxy: Connecting to ResourceManager at localhost/127.0.0.1:8032
18/06/19 19:35:46 WARN mapreduce.JobResourceUploader: Hadoop command-line option parsing not
performed. Implement the Tool interface and execute your application with ToolRunner to remedy this.
18/06/19 19:35:46 INFO input.FileInputFormat: Total input paths to process : 1
18/06/19 19:35:46 INFO mapreduce.JobSubmitter: number of splits:1
18/06/19 19:35:46 INFO mapreduce.JobSubmitter: Submitting tokens for job:
job_1529402911534_0004
18/06/19 19:35:46 INFO impl.YarnClientImpl: Submitted application application_1529402911534_0004
18/06/19 19:35:47 INFO mapreduce.Job: The url to track the job: [http://localhost:8088/proxy/](#)
application_1529402911534_0004/
18/06/19 19:35:47 INFO mapreduce.Job: Running job: job_1529402911534_0004
18/06/19 19:35:54 INFO mapreduce.Job: Job job_1529402911534_0004 running in uber mode : false
18/06/19 19:35:54 INFO mapreduce.Job: map 0% reduce 0%
18/06/19 19:36:00 INFO mapreduce.Job: map 100% reduce 0%
18/06/19 19:36:00 INFO mapreduce.Job: Job job_1529402911534_0004 completed successfully
18/06/19 19:36:00 INFO mapreduce.Job: Counters: 30
File System Counters
FILE: Number of bytes read=0
FILE: Number of bytes written=107391
FILE: Number of read operations=0
FILE: Number of large read operations=0
FILE: Number of write operations=0
HDFS: Number of bytes read=836
HDFS: Number of bytes written=685
HDFS: Number of read operations=5
HDFS: Number of large read operations=0
HDFS: Number of write operations=2
Job Counters
Launched map tasks=1
Data-local map tasks=1
Total time spent by all maps in occupied slots (ms)=4091
Total time spent by all reduces in occupied slots (ms)=0

Total time spent by all map tasks (ms)=4091
Total vcore-milliseconds taken by all map tasks=4091
Total megabyte-milliseconds taken by all map tasks=4189184
Map-Reduce Framework
Map input records=18
Map output records=16
Input split bytes=119
Spilled Records=0
Failed Shuffles=0
Merged Map outputs=0
GC time elapsed (ms)=45
CPU time spent (ms)=390
Physical memory (bytes) snapshot=130457600
Virtual memory (bytes) snapshot=2056757248
Total committed heap usage (bytes)=127795200
File Input Format Counters
Bytes Read=717
File Output Format Counters
Bytes Written=685
You have new mail in /var/spool/mail/acadgild
[acadgild@localhost ~]$ hdfs dfs -ls /hadoop1/mapreduce
18/06/19 19:36:15 WARN util.NativeCodeLoader: Unable to load native-hadoop library for your
platform... using builtin-java classes where applicable
Found 5 items
-rw-r--r-- 1 acadgild supergroup
29 2018-06-19 15:23 /hadoop1/mapreduce/1.txt
-rw-r--r-- 1 acadgild supergroup
717 2018-06-19 19:33 /hadoop1/mapreduce/television.txt
drwxr-xr-x - acadgild supergroup
0 2018-06-19 19:35 /hadoop1/mapreduce/tvinvalidout
drwxr-xr-x - acadgild supergroup
0 2018-06-19 15:39 /hadoop1/mapreduce/wordcountout
drwxr-xr-x - acadgild supergroup
0 2018-06-19 15:39 /hadoop1/mapreduce/wordout
[acadgild@localhost ~]$ hdfs dfs -ls /hadoop1/mapreduce/tvinvalidout
18/06/19 19:36:36 WARN util.NativeCodeLoader: Unable to load native-hadoop library for your
platform... using builtin-java classes where applicable
Found 2 items
-rw-r--r-- 1 acadgild supergroup
0 2018-06-19 19:35 /hadoop1/mapreduce/tvinvalidout/
_SUCCESS
-rw-r--r-- 1 acadgild supergroup
685 2018-06-19 19:35 /hadoop1/mapreduce/tvinvalidout/part-
m-00000
[acadgild@localhost ~]$ hdfs dfs -cat /hadoop1/mapreduce/tvinvalidout/part-m-00000
18/06/19 19:37:10 WARN util.NativeCodeLoader: Unable to load native-hadoop library for your
platform... using builtin-java classes where applicable
1
Samsung|Optima|14|Madhya Pradesh|132401|14200
2
Onida|Lucid|18|Uttar Pradesh|232401|16200
3
Akai|Decent|16|Kerala|922401|12200
4
Lava|Attention|20|Assam|454601|24200
5
Zen|Super|14|Maharashtra|619082|9200

6
Samsung|Optima|14|Madhya Pradesh|132401|14200
7
Onida|Lucid|18|Uttar Pradesh|232401|16200
8
Onida|Decent|14|Uttar Pradesh|232401|16200
9
Lava|Attention|20|Assam|454601|24200
10 Zen|Super|14|Maharashtra|619082|9200
11 Samsung|Optima|14|Madhya Pradesh|132401|14200
12 Samsung|Decent|16|Kerala|922401|12200
13 Lava|Attention|20|Assam|454601|24200
14
15
16
Samsung|Super|14|Maharashtra|619082|9200
Samsung|Super|14|Maharashtra|619082|9200
Samsung|Super|14|Maharashtra|619082|9200
Code

```
package com.acadgild.hadoop.mapreduce.tvsales;
import
import
import
import
import
import
import
import
java.io.IOException;
org.apache.hadoop.conf.Configuration;
org.apache.hadoop.fs.Path;
org.apache.hadoop.io.IntWritable;
org.apache.hadoop.io.Text;
org.apache.hadoop.mapreduce.Job;
org.apache.hadoop.mapreduce.Mapper;
org.apache.hadoop.mapreduce.lib.input.FileInputFormat;
org.apache.hadoop.mapreduce.lib.output.FileOutputFormat;
public class invalidtv {
public static class invalidtvMapper
extends Mapper<Object, Text,IntWritable, Text> {
Text validtv;
IntWritable count;
public void setup(Context context) {
validtv= new Text();
count=new IntWritable();
}
/** filtering invalid sales*/
public void map(Object key, Text value, Context context
) throws IOException, InterruptedException {
String[] lineArray = value.toString().split("\\|");
if (!(lineArray[0].matches("NA")) && !(lineArray[1].matches("NA"))) {
validtv.set(value);
count.set(count.get()+1);
//print the count of valid tv
context.write(count, validtv);
}
```

```
}
}
public static void main(String[] args) throws Exception {
Configuration conf = new Configuration();
Job job = Job.getInstance(conf, "tv invalidrecords");
job.setJarByClass(invalidtv.class);
job.setMapOutputKeyClass(IntWritable.class);
job.setMapOutputValueClass(Text.class);
job.setOutputKeyClass(IntWritable.class);
job.setOutputValueClass(Text.class);
job.setMapperClass(invalidtvMapper.class);
job.setNumReduceTasks(0);
FileInputFormat.addInputPath(job, new Path(args[0]));
FileOutputFormat.setOutputPath(job, new Path(args[1]));
job.waitForCompletion(true);
}
}
```

Task2
code

```
package com.acadgild.hadoop.mapreduce.tvcompany;

import java.io.IOException;
import org.apache.hadoop.conf.Configuration;
import org.apache.hadoop.fs.Path;
import org.apache.hadoop.io.IntWritable;
import org.apache.hadoop.io.Text;
import org.apache.hadoop.mapreduce.Job;
import org.apache.hadoop.mapreduce.Mapper;
import org.apache.hadoop.mapreduce.Reducer;
import org.apache.hadoop.mapreduce.lib.input.FileInputFormat;
import org.apache.hadoop.mapreduce.lib.output.FileOutputFormat;

public class tvcompany {
public static class tvcompanyMapper
extends Mapper<Object, Text, Text,IntWritable> {
Text companyname;
IntWritable count;
public void setup(Context context) {
companyname= new Text();
count=new IntWritable(1);
}
/** filtering invalid sales*/
public void map(Object key, Text value, Context context
) throws IOException, InterruptedException {
String[] lineArray = value.toString().split("\\|");
if (!(lineArray[0].matches("NA")) && !(lineArray[1].matches("NA"))) {
companyname.set(lineArray[0]);
context.write(companyname,count);
            }
        }
    }
public static class tvcompanyReducer
extends Reducer<Text,IntWritable,Text,IntWritable> {
public void reduce(Text key, Iterable<IntWritable> values,
```

```
            Context context) throws IOException, InterruptedException {
int sum = 0;
for (IntWritable val : values) {
 sum += val.get();
            }
context.write(key, new IntWritable(sum));
       }
}
public static void main(String[] args) throws Exception {
Configuration conf = new Configuration();
Job job = Job.getInstance(conf, "tv sales for each company");
job.setJarByClass(tvcompany.class);

job.setMapOutputKeyClass(Text.class);
job.setMapOutputValueClass(IntWritable.class);

job.setOutputKeyClass(Text.class);
job.setOutputValueClass(IntWritable.class);

job.setMapperClass(tvcompanyMapper.class);
job.setReducerClass(tvcompanyReducer.class);
job.setCombinerClass(tvcompanyReducer.class);
job.setNumReduceTasks(1);

FileInputFormat.addInputPath(job, new Path(args[0]));
FileOutputFormat.setOutputPath(job, new Path(args[1]));
job.waitForCompletion(true);
}
}
```

[acadgild@localhost ~]$ hdfs dfs -cat /hadoop1/mapreduce/tvcompanyout-t/part-r-00000
18/06/20 11:48:58 WARN util.NativeCodeLoader: Unable to load native-hadoop library for your
platform... using builtin-java classes where applicable
Akai 1
Lava   3
Onida   3
Samsung 7
Zen 2

Task3
code

```
package com.acadgild.hadoop.mapreduce.tvcompany;

import java.io.IOException;
import org.apache.hadoop.conf.Configuration;
import org.apache.hadoop.fs.Path;
import org.apache.hadoop.io.IntWritable;
import org.apache.hadoop.io.Text;
import org.apache.hadoop.mapreduce.Job;
import org.apache.hadoop.mapreduce.Mapper;
import org.apache.hadoop.mapreduce.Reducer;
import org.apache.hadoop.mapreduce.lib.input.FileInputFormat;
import org.apache.hadoop.mapreduce.lib.output.FileOutputFormat;

public class onidasale {
public static class onidasaleMapper
```

```
extends Mapper<Object, Text, Text,IntWritable> {
Text statename;
IntWritable count;
public void setup(Context context) {
statename= new Text();
count=new IntWritable(1);
}
/** filtering invalid sales and storing the state names saling onida*/
public void map(Object key, Text value, Context context
) throws IOException, InterruptedException {
String[] lineArray = value.toString().split("\\|");
if (!(lineArray[0].matches("NA")) && !(lineArray[1].matches("NA"))) {
statename.set(lineArray[3]);
context.write(statename,count);
                }
            }
        }
public static class onidasaleReducer
extends Reducer<Text,IntWritable,Text,IntWritable> {
public void reduce(Text key, Iterable<IntWritable> values,
            Context context) throws IOException, InterruptedException {
int sum = 0;
for (IntWritable val : values) {
 sum += val.get();
            }
context.write(key, new IntWritable(sum));
        }
}
public static void main(String[] args) throws Exception {
Configuration conf = new Configuration();
Job job = Job.getInstance(conf, "onida tv salling states");
job.setJarByClass(onidasale.class);

job.setMapOutputKeyClass(Text.class);
job.setMapOutputValueClass(IntWritable.class);

job.setOutputKeyClass(Text.class);
job.setOutputValueClass(IntWritable.class);

job.setMapperClass(onidasaleMapper.class);
job.setReducerClass(onidasaleReducer.class);
job.setNumReduceTasks(1);

FileInputFormat.addInputPath(job, new Path(args[0]));
FileOutputFormat.setOutputPath(job, new Path(args[1]));
job.waitForCompletion(true);
}
}
```

Ans
[acadgild@localhost ~]$ hdfs dfs -cat /hadoop1/mapreduce/onidasaleout/part-r-0000018/06/20
12:13:24 WARN util.NativeCodeLoader: Unable to load native-hadoop library for your platform... using
builtin-java classes where applicable
Assam   3
Kerala   2
Madhya Pradesh   3
Maharashtra   5